Lexical Semantic Change: Models, Data and Evaluation

LREC 2022 - Tutorial - 20 June 2022

Pierpaolo Basile[1], Annalina Caputo[2], Pierluigi Cassotti[1] and Rossella Varvara[3]
University of Bari[1], Dublin City University[2], Université de Fribourg[3]

Outline

- Introduction
- Data: diachronic corpora and diachronic resources; annotation of lexical semantic change
- Models: background concepts, alignment models, contextual models
- Evaluation: tasks (binary, ranking, temporal alignment, LSC discovery), results

Introduction

Semantic change

Lexical semantic Change (LSC): diachronic evolution of the meaning of a word.
e.g. Italian *positivo* (“positive”) meaning “tested positive to covid-19”

≠ neologisms (e.g. *quarantini,* cocktail consumed in quarantene)

≠ Grammatical semantic change and grammaticalization
e.g. English ‘will’: from ‘want, desire’ (ex. I will more silver) to future tense (ex. It will rain tomorrow)

Types of lexical semantic change

- **pejoration**: changes whose result is a more negative meaning
- **amelioration**: changes whose result is a more positive meaning

Ex. *rude* from ‘unmannered’ to ‘physically attractive’ (examples from Hollmann 2009)

> *I didn't wanna bowl over all geezer and rude, Not rude as in good but just rude like uncouth*

Types of lexical semantic change

- **pejoration**: changes whose result is a more negative meaning
- **amelioration**: changes whose result is a more positive meaning

e.g. *silly* 'foolish, stupid' < ME *sely* 'happy, innocent' < OE *sælig* ''blessed, blissful'

Types of lexical semantic change

- **broadening** (generalization or extension or widening)
- **narrowing** (restriction or specialization)

Ex. *dog*: from a more specific meaning in Old English (``dog of a powerful breed'') to the more general term (Traugott 2006)

Ex. *deor - deer:* from the larger meaning of ‘animal’, to the narrower reference in present-day English.

Types of processes of semantic change

Semantic changes can be further classified on the basis of the cognitive process that originated them, e.g. from **metonymy** or **metaphor**.

e.g. metonymic shift of *rude*, vulgarity as part of attractiveness

Types of originating factors

- language-internal factors
- language-external factors

e.g. *cell*: from ``prisoner cell'' to ``cell phone'' is a case of change due to language-external factors (Hamilton et al. 2016).

Regularities in lexical semantic change

Usually words change from relatively objective meanings into increasingly subjective ones (**subjectification**, Traugott 1989).

Why automatic LSC detection?

- Related to word sense disambiguation
- New insights for the study of historical linguistics
- Automatic classification of types of semantic shift
- More accurate models of word meaning that consider temporal differences

Growing interest in the community

- 3rd International Workshop on Computational Approaches to Historical Language Change (https://languagechange.org/events/2022-acl-lchange/)
- Project “Towards Computational Lexical Semantic Change Detection” (https://languagechange.org/)
- SemEval 2020 Task 1 - Unsupervised Lexical Semantic Change Detection (Schlechtweg et al. 2020)
- DIACR-Ita task at the EVALITA 2020 (Basile et al. 2020)
- RuShiftEval (Kutuzov et al. 2021)